ARTICLE IN PRESS

Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Engineering Management—Article

Research on Active Safety Methodologies for Intelligent Railway Systems

Yong Qin [a,c,*], Zhiwei Cao [a,c], Yongfu Sun [b], Linlin Kou [d], Xuejun Zhao [e], Yunpeng Wu [f], Qinghong Liu [g], Mingming Wang [g], Limin Jia [a,*]

[a] State Key Laboratory of Advanced Rail Autonomous Operation, Beijing Jiaotong University, Beijing 100044, China
[b] Chinese Academy of Engineering, Beijing 100088, China
[c] Key Laboratory of Railway Industry of Proactive Safety and Risk Control, Beijing Jiaotong University, Beijing 100044, China
[d] Beijing Subway Operation Technology Centre, Beijing 100044, China
[e] CRRC Academy Co., Ltd., Beijing 100070, China
[f] School of Safety Engineering and Emergency Management, Shijiazhuang Tiedao University, Shijiazhuang 050043, China
[g] China Academy of Railway Sciences, Beijing 100081, China

ARTICLE INFO

Article history:
Received 12 April 2022
Revised 7 June 2022
Accepted 27 June 2022
Available online 1 November 2022

Keywords:
Intelligent railway system
Active safety methodology
Prognostics and health management
Intelligent surrounding perception
Operation and maintenance risk control

ABSTRACT

Safety is essential when building a strong transportation system. As a key development direction in the global railway system, the intelligent railway has safety at its core, making safety a top priority while pursuing the goals of efficiency, convenience, economy, and environmental friendliness. This paper describes the state of the art and proposes a system architecture for intelligent railway systems. It also focuses on the development of railway safety technology at home and abroad, and proposes the active safety method and technology system based on advanced theoretical methods such as the in-depth integration of cyber-physical systems (CPS), data-driven models, and intelligent computing. Finally, several typical applications are demonstrated to verify the advancement and feasibility of active safety technology in intelligent railway systems.

1. Introduction

With the continuous growth of China's economy, its rail transit systems, such as high-speed railways and urban rail transit, are developing rapidly. As of the end of 2021, China had 150 000 km of railways in operation, 40 000 km of which are high-speed railways, making China the country with the longest high-speed railway in the world. In terms of urban rail transit, 50 cities in China have a total of more than 9 192 km of operating routes, with an operating mileage that is also ranked first in the world.

Rail transit is the backbone of China's transportation system and its main force in realizing the national development strategy of "building a strong transportation country" [1], which proposes the construction of a safe, convenient, efficient, green, and economic transportation system. Safety is the first goal of this strategy. The strategy also proposes speeding up the construction of new transportation infrastructure. The intelligent railway system is the key to technological innovation in rail transit in the future, and safety is its first consideration.

Safety is regarded as the core competitiveness of rail transit systems around the world; the safety management level of China's railway system has also reached the international first-class level. However, as mentioned in Ref. [1], there is a need for high-quality development, the scale and density of system operations are constantly expanding, and the system complexity of engineering technology is increasing. The safety of China's rail transit system presents severe challenges, and there is an urgent need for research on the development mode, system framework, and key technologies.

In this paper, we propose an active safety framework for intelligent railway systems and demonstrate typical applications of active safety methodologies in railways. The main contributions of this paper are as follows. First, it summarizes and describes the development stages and key technologies of intelligent railways and proposes a system architecture for the intelligent railway. Second, it proposes an active safety system according to the development of railway safety technology at home and abroad, and presents a closed-loop model and hierarchical structure at

* Corresponding authors.
E-mail addresses: yqin@bjtu.edu.cn (Y. Qin), lmjia@bjtu.edu.cn (L. Jia).

https://doi.org/10.1016/j.eng.2022.06.025
2095-8099/

Please cite this article as: Y. Qin, Z. Cao, Y. Sun et al., Research on Active Safety Methodologies for Intelligent Railway Systems, Engineering, https://doi.org/10.1016/j.eng.2022.06.025

the logical and physical levels, respectively. Finally, it demonstrates several typical applications to verify the advancement and feasibility of active safety technology in the intelligent railway.

The remainder of this paper is organized as follows. Section 2 summarizes the future direction of intelligent railway systems. On this basis, Section 3 focuses on the development trend of railway system safety technology at home and abroad, and proposes the active safety method and technology system. Section 4 demonstrates the advancement and feasibility of active safety technology in intelligent railways through several typical applications. Finally, conclusions are drawn in Section 5.

2. The intelligent railway system 2.0

To realize a safer, more efficient, more comfortable, and greener railway transportation system, the concept of the intelligent railway system 2.0 is utilized. This concept makes full use of advanced information and communication, artificial intelligence, the Internet of Things (IoT), big data, robots, and other advanced technologies, and takes self-perception, self-learning, self-decision-making, and self-control as its core processes. The intelligent railway system 2.0 can provide efficient, accurate, and personalized displacement services based on the optimal management and control of equipment and facilities [2].

2.1. An overview of intelligent railway development

The development of the smart railway has been a gradual process with three milestone stages: the digital railway, the intelligent railway 1.0, and the intelligent railway 2.0. Fig. 1 describes and contrasts the development stages and key technologies of the intelligent railway in detail. The digital railway improved railway informatization by transforming all resources of the railway and its operating environment into digital computing resources. The intelligent railway 1.0 realized the network connection and collaborative optimization of transportation services by using information networks, interconnection, and other technologies to integrate the scattered data of different specialties and different businesses into unified information in time and space. The intelligent railway 2.0 organizes, analyzes, and processes resources using technologies such as big data, the IoT, and artificial intelligence in order to realize autonomous and intelligent perception, learning, decision-making, and execution. In this way, it provides comfortable and convenient passenger and freight services economically, efficiently, and safely. In the development of the intelligent railway, the application of emerging technologies has played a pivotal role. Key technologies such as the advanced IoT, cloud computing, the mobile Internet, big data, artificial intelligence, and new materials are having an enormously transformative effect on the internal functions and external forms of the railway system.

To maintain and improve the competitiveness of rail transit in the 21st century, several countries with developed railway systems have taken the intelligent railway as an important development direction. The European Union (EU) has proposed a development strategy for a safer, greener, and smarter European rail transit system and had formulated development plans such as "Rail route 2050: towards a competitive, resource-efficient, and intelligent rail transport system for 2050," "Rail 2050 vision rail—the backbone of Europe's mobility," and "Shift2Rail" [3–5]. The United Kingdom has proposed the "Great British railways: the Williams–Shapps plan for rail," in which the current challenges presented by the British railway system are considered and research on the intelligent railway is regarded as an important task [6]. The United States has proposed the initiative "Beyond traffic 2045" [7], and IBM has released "Think beyond the rails: leading in 2025" to promote the technological development of the railway industry [8]. Japan has proposed several development plans, such as the "Grand design of national spatial development towards 2050" and "Research 2025" [9,10].

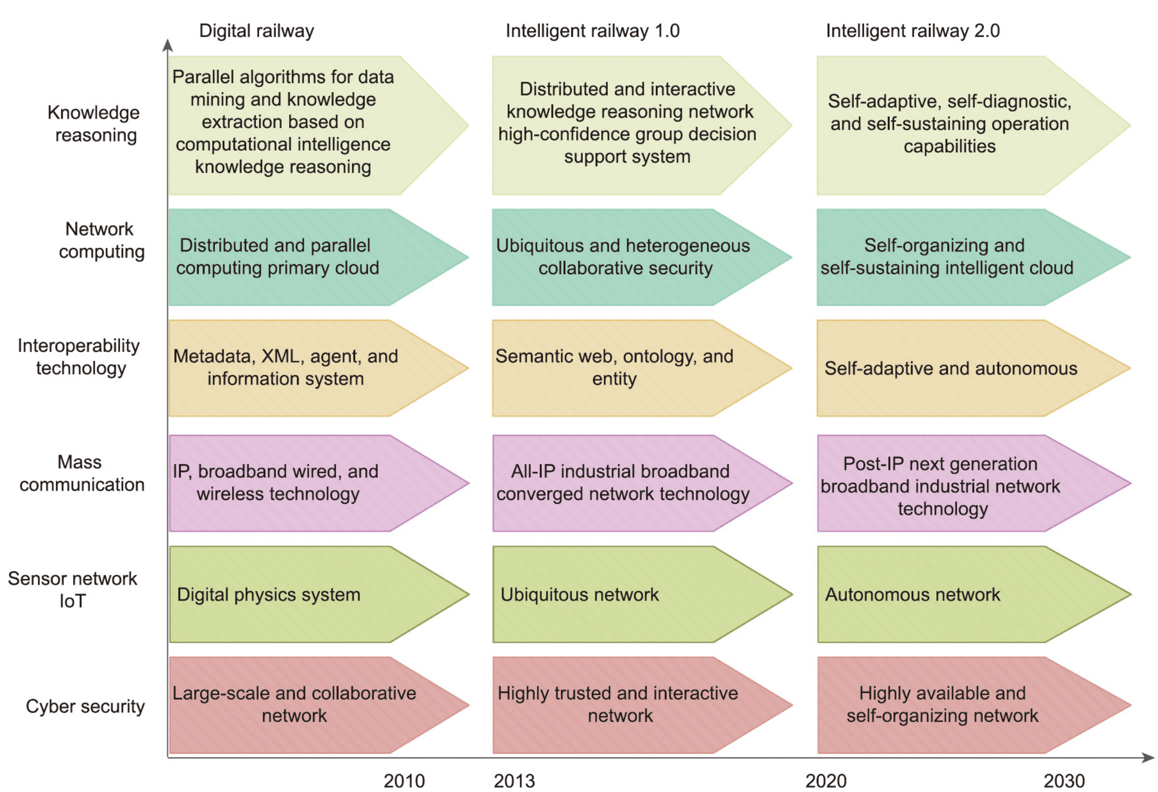

Fig. 1. Development stages and key technologies of the intelligent railway. XML: eXtensible Markup Language; IP: Internet protocol.

2.2. Intelligent railway system framework

An intelligent railway system framework describes the main functions of the system and their interrelations, including the logical framework and the physical framework. The *logical framework* depicts the functional modules and the interaction between the modules, and works as the top-level design framework of the intelligent railway system. As shown in Fig. 2, it includes the physical layer, perception layer, information fusion layer, intelligent analysis layer, business optimization layer, collaborative service layer, and system goal layer.

An intelligent railway system is a complex physical system consisting of multiple onsite subsystems. As shown in Fig. 3, the *physical framework* of the intelligent railway 2.0 realizes mutual functional support based on ubiquitous perception, an information integration platform, and comprehensive intelligent decision-making. The physical subsystems include intelligent trains, intelligent lines, intelligent stations, intelligent yards, intelligent security, intelligent transportation management, intelligent service, and an information integration platform.

The core process of the intelligent railway 2.0 is a closed-loop iterative process of self-perception, self-learning, self-decision-making, and self-control, which is highly autonomous. *Self-perception* acquires ubiquitous information on the operating state of the railway system; *self-learning* transforms this acquired information into useful knowledge for modeling and reasoning; *self-decision-making* performs scientific reasoning based on the knowledge model; and *self-control* outputs highly safe and automated execution. Therefore, the intelligence level of the railway system is continuously improved through a long-term iterative process. In the intelligent railway 2.0, active safety plays the role of top-level design. Self-perception, self-learning, self-decision-making, and self-control all work under the guidance of active safety, which makes the railway system increasingly automated and intelligent. From the perspective of complex systems and informatics, the intelligent railway is a typical dispersion structure, which is in a self-organized state that is far from a balanced state. Active safety causes the intelligent railway to constantly transform information data into knowledge and decision-making, thereby generating negative entropy to reduce the system entropy and form an orderly structure.

3. Active safety methodology

The concept of active safety is rapidly emerging with the safe operation of complex and large-scale system engineering and the development of IoT technology. Compared with passive safety, active safety is based on system safety theory and emphasizes the real-time perception and control output of the system, thereby reducing system risks in advance and avoiding accidents. Active safety methodology is a theory that models, analyzes, and optimizes the dynamic relationship of risk factors in complex systems. In recent years, active safety has become an interdisciplinary research field connecting safety science, control science, information science, system science, and other disciplines, while playing an increasingly important role in transportation, electric power, manufacturing, the Internet, the military, and other fields.

3.1. Development trend of active safety technology in railways

Research on active safety technology is mainly concentrated in Europe, the United States, China, and Japan. The research in these regions and countries is very similar in terms of key technologies, but the specific areas of application are different. Table 1 introduces the development trends of active safety technology in these regions and countries.

In "Rail 2050 vision rail—the backbone of Europe's mobility" [4], Europe proposes a safety goal of "zero fatalities" for the Europe railway system by 2050. To achieve this goal, continuous technological innovation is required, such as: predictive maintenance of infrastructure based on technologies such as the IoT, machine learning, artificial intelligence, and big data analysis; fully automated onsite maintenance and repair based on robots; the safe operation of manned and unmanned smart trains on the same railway network; and wide use of the virtual simulation safety certification technology of rolling stock based on digital modeling to reduce physical entity testing.

"Beyond traffic 2045" [7] proposes that US railway companies and railway authorities should continue to focus on safety, build a railway safety culture, and develop and implement new safety technologies to drive future safety improvements. More specifically, these new technologies should include tracking health monitoring technology, online monitoring of train equipment performance and condition repair technology, and unmanned aerial vehicle (UAV) research.

China proposes building itself into a modern and strong railway system with safe and high-quality services that are strongly guaranteed and with world-leading strength by 2035 [11]. China is establishing a safety system of personal protection, material protection, and technical protection, and is comprehensively improving the safety level, safety prevention, management and control capabilities, and emergency response and rescue capabilities of its railway system, so that railway safety can be reliably ensured and the accident and death rate of China's railway traffic can be significantly reduced.

The "Grand design of national spatial development towards 2050" [9] describes the strategic direction of Japan's railway development, which relies on advanced technologies such as big data, the Internet, and autonomous driving to build a safer railway system. Japan is paying special attention to research that can help improve the capacity of railways to cope with natural disasters, such as heavy rains, strong winds, and large earthquakes, which are frequent and increasingly severe [10].

It can be seen from the development trends described above that active safety has become a universally recognized method and mainstream research direction for the future railway safety needs of all countries. All countries with a strong railway system are developing core technologies and equipment systems to ensure that their railway safety level is maintained at a first-class global level.

3.2. Active safety system

As shown in Fig. 4, an active safety system is an iteratively optimized closed-loop model composed of core technologies and typical technical processing steps. The model takes as its research objects the security elements of a complex system that includes people, equipment, the environment, and management, and forms the four closely related technical processing steps of *safety monitoring*, *risk assessment*, *risk control*, and *emergency response*. As shown in Fig. 4, these processing steps are based on core technologies such as obtaining status, fault diagnosis, feature analysis, operation and maintenance optimization, and contingency planning. The closed-loop iterative optimization of the four processing steps continuously reduces the system risk level and ensures the safe operation of the system. This closed-loop model has the following technical characteristics:

(1) **Data-driven basis:** The model includes automatic identification and intelligent knowledge-based reasoning according to real-time monitoring data.

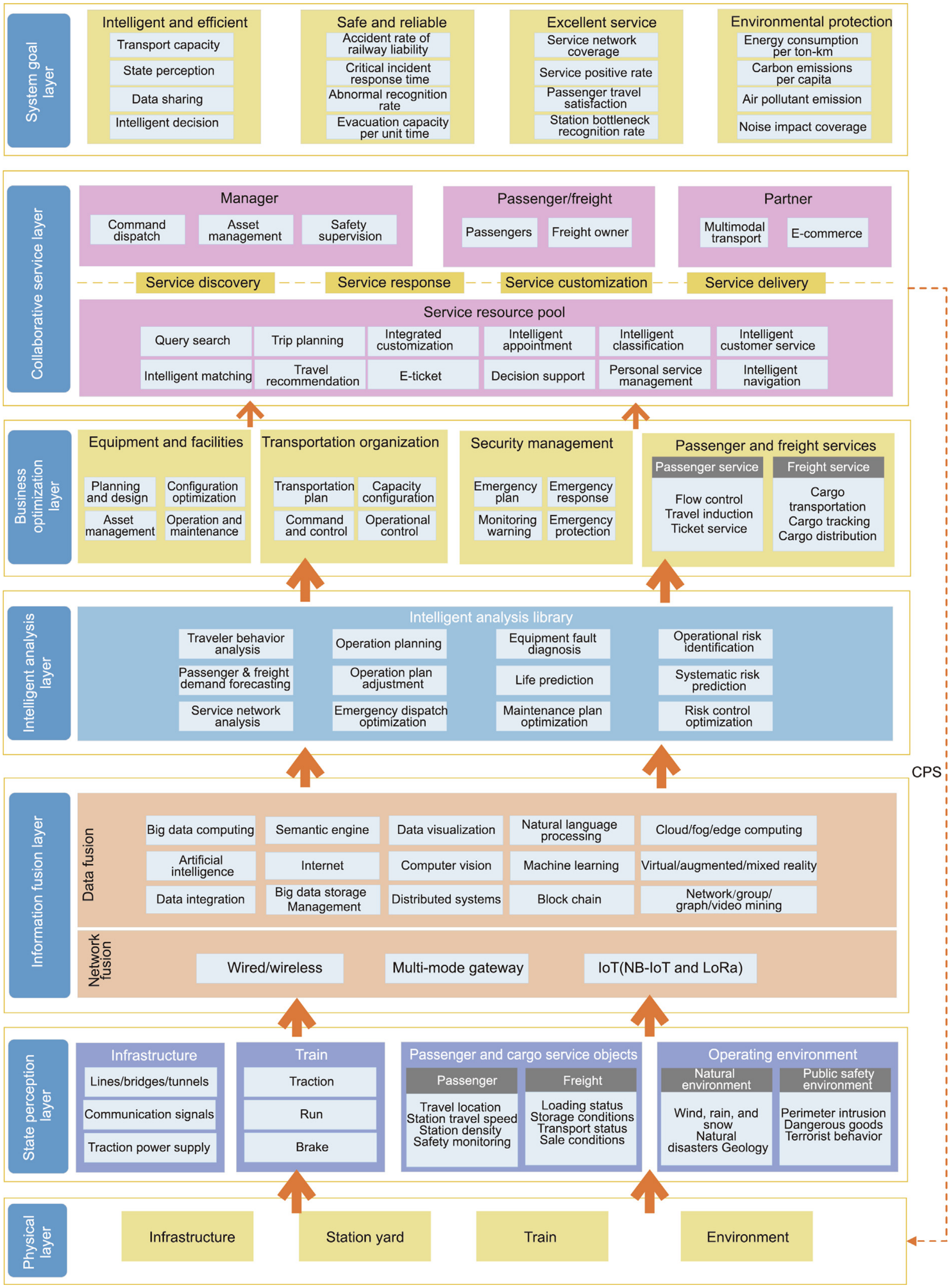

Fig. 2. The logical framework of the intelligent railway 2.0. CPS: cyber-physical systems; NB-IoT: narrow band Internet of Things; LoRa: long range radio.

(2) **Global dynamic risk management and control:** The model covers the whole process from risk identification to risk control, from a system perspective.

(3) **Advanced information-processing technology:** The model makes full use of advanced technologies such as the IoT, sensors, and big data.

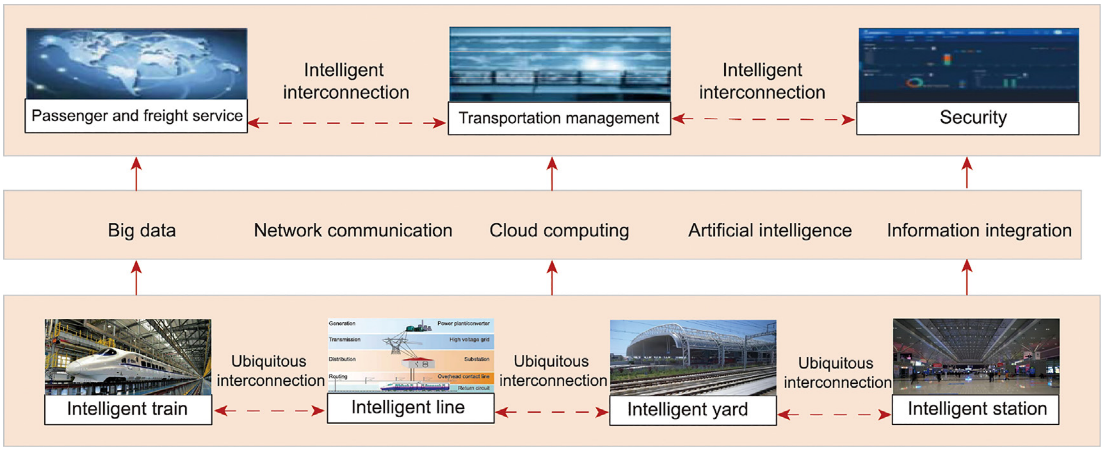

Fig. 3. The physical framework of the intelligent railway 2.0.

Table 1
The development of active safety technology in Europe, the United States, China, and Japan.

Country/ region	Key technologies	Applications
Europe	IoT, machine learning, artificial intelligence and big data, robots, unmanned technology, and virtual simulation	Predictive maintenance of infrastructure, fully automated onsite maintenance and repair, unmanned smart trains
United	States	IoT, artificial intelligence, big data, robots, unmanned technology, and UAVs
Tracking	health monitoring, online monitoring of trains, railway inspections, automation, and connected vehicles	
China	IoT, 5G, big data, artificial intelligence, robots, unmanned technology, UAVs, and BeiDou Navigation Satellite System	Self-awareness of operating status, self-diagnosis of equipment failure, self-decision-making for guiding safety, and an intelligent monitoring system integrating air, earth, and vehicles
Japan	IoT, 5G, big data, artificial intelligence, robots, unmanned technology, and virtual simulation	Natural disaster prediction and detection, autonomous train operation and control, digital maintenance

5G: the fifth generation mobile communication technology.

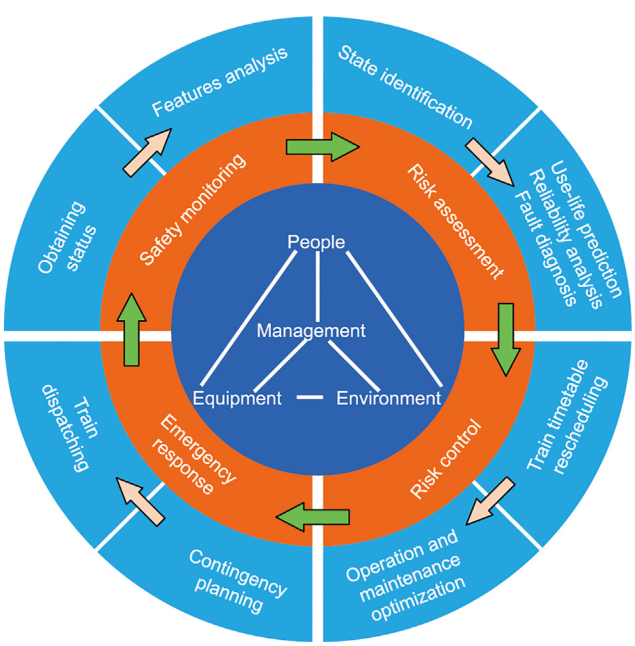

Fig. 4. A closed-loop model of the active safety system.

As shown in Fig. 5, the active safety system has a typical hierarchical structure on a physical framework, which includes a perception layer, intelligence layer, and system layer. The *perception layer* uses advanced sensors and communication technologies for real-time monitoring and deep information fusion. The intelligence layer performs fault diagnosis, risk assessment, reliability calculation, and prediction based on the feature extraction and state identification of risk-state data. The *system layer* provides integrated monitoring, evaluation, regulation, emergency response, and other system functions to users at different levels and task types in order to achieve system-level risk management and control.

4. Typical application of active safety technology

In recent years, active safety technology in railway systems has improved significantly. In this section, we introduce some of the research achievements and progress made by our research team. In these studies, trains, the train operating environment, and the railway network are regarded as research objects, forming a research framework from local to global. First, a train is perceived as a complex system with a large number of devices; thus, it is necessary to use prognostics and health management technology to ensure the safety of the train system and key equipment. Next, the operating environment of the train is perceived as an open environment that often involves unpredictable hazards, so intelligent perception methods are proposed to solve these problems. Finally, to improve the operating efficiency and reduce economic losses after accidents, we study operation risk assessment and the intelligent scheduling of railway networks.

4.1. The key technology of prognostics and health management for a railway train

4.1.1. A weak fault diagnosis method based on signal reconstruction and entropy

The health status of key equipment is critical to safety during train operation. The locomotives and electrical multiple units in

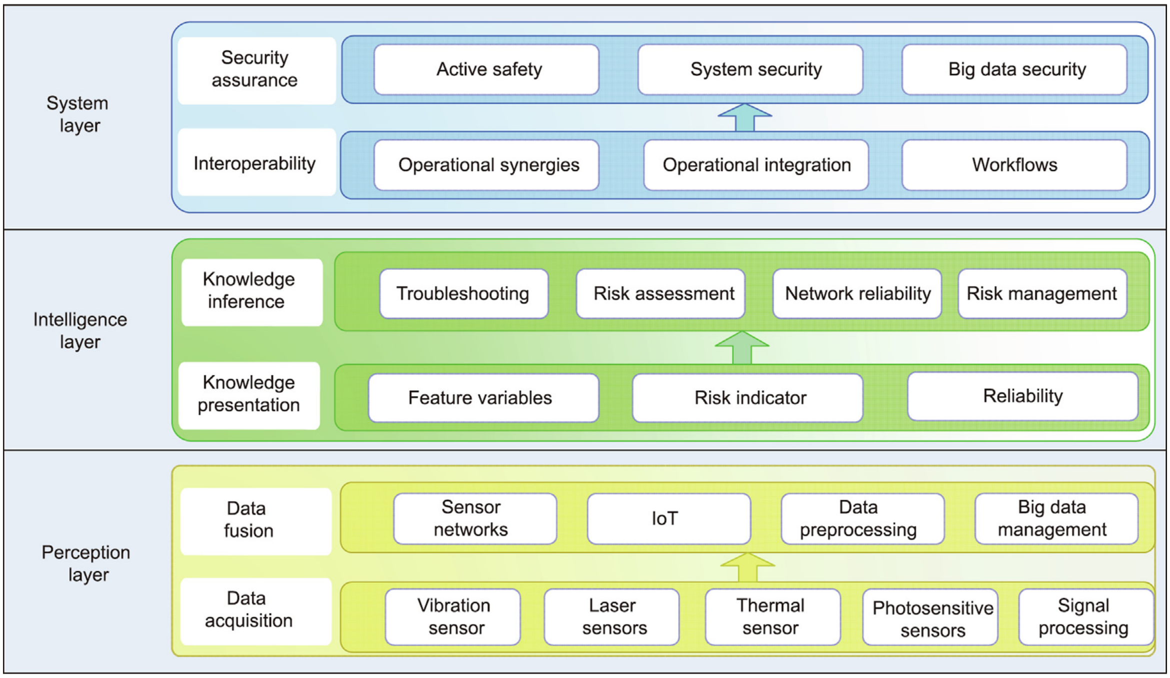

Fig. 5. The hierarchical structure of the active safety system.

Chinese railway systems have gradually been installed with sensors to monitor the status of the equipment in real time. However, the number of sensors is limited by the installing environment; as a result, it is generally not possible to meet the requirement for the number of observed signals to be less than the number of source signals in the determined blind source separation algorithm. For example, wheelset bearing observation signals are usually single-channel signals, so it is necessary to use single-channel signals to extract fault feature blind sources—that is, for the signal reconstruction of underdetermined blind sources in the fault feature. In Ref. [12], we propose an underdetermined blind source extraction method for the weak fault features of wheelset bearings based on signal decomposition and kernel correlation theory. First, the empirical mode decomposition algorithm decomposes the observed signals into several virtual channels and transforms the underdetermined blind source extraction into determined blind source extraction. Next, the number of signal sources is estimated based on a Bayesian selection model. Finally, the objective function is maximized based on the correlation of signals in the regenerated Hilbert space, a fixed-point algorithm is used to solve the problem, and the extracted vector and weak fault characteristics of the wheelset bearings are obtained.

Based on a reconstruction of the fault characteristic signals, Ref. [13] proposes an intelligent fault diagnosis method for wheelset bearings based on cyclic correlation entropy and narrowband filtering, as shown in Fig. S1 in Appendix A. First, the fault signal features are extracted based on cyclic correlation entropy and narrowband filtering. Then, a least-squares support vector machine is used to train the extracted feature samples, and the test sample set is classified to realize intelligent fault diagnosis of the wheelset bearings.

To verify the effectiveness of the method, the algorithm is analyzed using a faulty bearing dataset obtained from the train wheelset bearing test bed. The selected fault modes are: no fault (NF), rolling element fault (REF), inner race fault (IRF), and outer race fault (ORF). The failure modes of the three groups of experiments are shown in Fig. S2 in Appendix A, in which NF, REF, IRF, and ORF are represented by blue, pink, black, and red dots, respectively. In Fig. S2(a), three samples of the ORF are misclassified, while the other fault modes are classified correctly. In Figs. S2(b) and (c), the method also performs well. These findings demonstrate the effectiveness of this model for bearing fault diagnosis under impact noise interference.

4.1.2. An end-to-end fault diagnosis method based on a multidimensional convolutional neural network (CNN)

The traditional equipment state recognition algorithm depends on the engineer's experience with the design characteristics, and it is difficult to ensure the quality of the results. Moreover, it is difficult to effectively fuse and utilize the multi-source information provided by multiple sensors. A multi-source information fusion method based on a CNN effectively avoids the complexity of artificial design features. As shown in Fig. 6, Ref. [14] proposes a fault diagnosis method based on a multidimensional CNN, which provides the necessary technical means for component performance warning and reliability evaluation. This method integrates the data obtained from multiple information sources, comprehensively evaluates the component state information in a timely manner, and realizes end-to-end component state identification. First, the multi-source signal is normalized to reduce noise interference and speed up program convergence. Then, the multi-source signals are fused in the data layer by using tensor expression. Finally, the sample data is trained with a CNN to obtain the optimal tensor domain, and the samples to be tested are mapped to the tensor domain in order to achieve automatic identification of the components.

To verify the effectiveness of the method, bogie vibration data from a Chinese railway system are used to analyze the algorithm. The dataset includes four bogie conditions: operation with no trouble, wheel out-of-roundness or flat, shaft misalignment, and wheel run out. In the experiment, the average accuracy is obtained by fivefold cross validation, the maximum number of iterations is 100, and the learning rate is 0.001. The results are shown in Fig. S3 in Appendix A; the accuracy is 100% in the training set and 99.16% in the testing set. These results demonstrate that the method is effective for bogie fault diagnosis in practice.

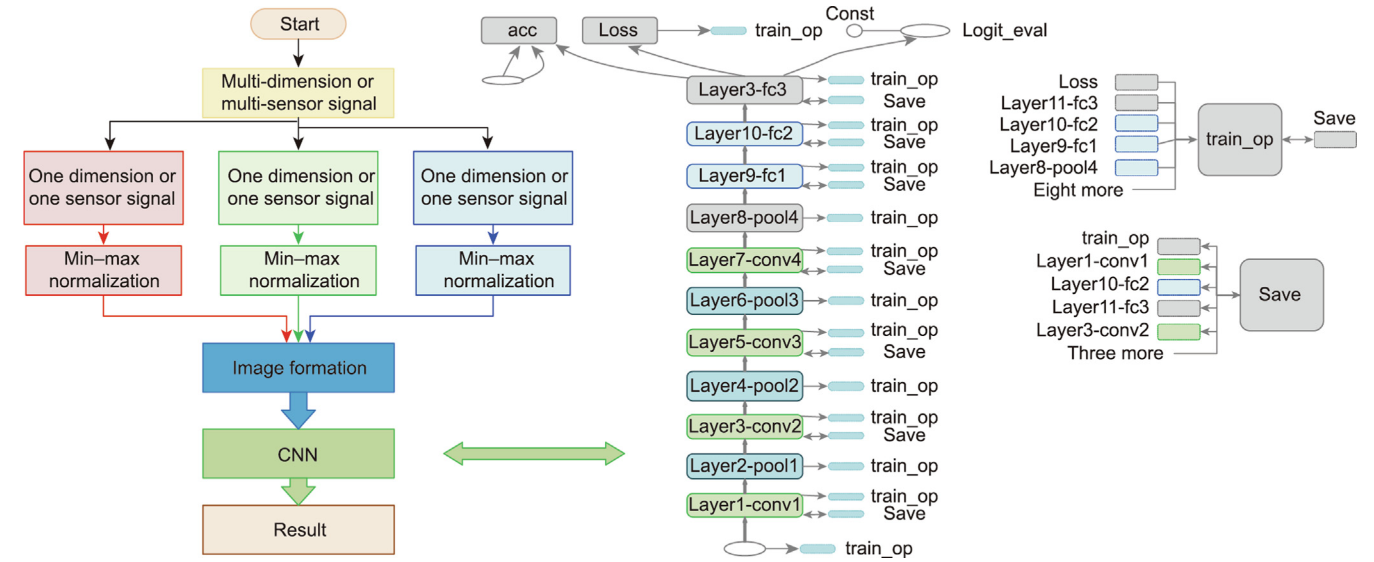

Fig. 6. The multidimensional end-to-end CNN model. acc: accuracy; train_op: train output; Const: constraint; fc: fully connected; logit_eval: logit evaluation; conv: convolution.

.

4.1.3. Life prediction of key train components based on a tensor domain and time-varying Markov process

In actual operation, some key train components must be replaced when they reach a certain number of kilometers. For the life prediction and reliability assessment of key train components, Ref. [15] proposes a model based on a fuzzy time sequence analysis method and a time-varying Markov chain. First, a fuzzy segmentation algorithm is used to automatically divide the component degradation sample data into multiple stages in order to match the scientific definition of the classification plane of the tensor domain. A dynamic time-warping algorithm regularizes the collected samples, provides effective data for studying degradation laws at different stages, eliminates the mapping between explicit and implicit states, and effectively reduces the diffusion of influence and trust. The segmented objective function and constraint conditions of a sample sequence are shown in Eqs. (1) and (2). Then, a state transition matrix related to the state residence time is introduced based on the semi-Markov process, and a time-varying transition probability matrix is proposed to describe the variation trend of the component characteristics with time. This provides a necessary means for a more accurate evaluation of the component state reliability and life prediction.

The objective function is

$$\mathrm{cost}(S_X^c) = \sum_{k=1}^{c}\sum_{i=1}^{N}\left(\mu_{k,i}\right)^m D^2(\mathbf{z}_i, \eta_k) \tag{1}$$

The constraint conditions are

$$\begin{aligned} &U = [\mu_{k,i}]_{c\times N},\ \ \mu_{k,i} \in [0,1], \forall k, i \\ &\sum_{k=1}^{c}\mu_{k,i} = 1,\ \ \forall i \\ &0 < \sum_{i=1}^{N}\mu_{k,i} < N,\ \ \forall k \end{aligned} \tag{2}$$

where the segmented objective function of the sample sequence is $\mathrm{cost}(S_X^c)$, X is time sequence, N is the data length of the time sequence, and $\mu_{k,i}$ is the membership degree of sample data point $\mathbf{z}_i$ in the kth tensor domain, with $k = 1, 2, \ldots, c$. In addition, $m = [1, \infty)$ is the fuzzy clustering weighted index; in general, $m = 2$. Moreover, $D^2(\mathbf{z}_i, \eta_k)$ is the distance between the sample data point and the clustering prototype, and η_k is the clustering prototype function of the kth tensor domain, which is a multivariate mixed Gaussian function.

The rolling bearing is one of the most widely used general mechanical parts in rail transit vehicles, and its failure rate is high. Machine abnormalities often occur, resulting in failure to work normally due to fatigue, wear, strain, electrical corrosion, fracture, bonding, and other reasons. Therefore, the bearing is taken as an example to verify the proposed method. For this purpose, whole-life bearing data collected on the PRONOSTIA experimental platform designed by the French Femto-ST Institute are adopted. The prediction results for the bearing named "Bearing1_1" are shown in Fig. S4 in Appendix A. The lifetime of the hidden Markov model (HMM) algorithm shows a state of stepwise decline, which is seriously inconsistent with the actual situation. The proposed algorithm is more accurate than the HMM in predicting life, as it accords with the trend of life decreasing with time. The algorithm adopts the tensor domain level of the current moment and the updated state transition probability matrix at this level, so the accuracy of the algorithm will become increasingly accurate with the gradual degradation of component performance.

4.1.4. Dynamic reliability analysis of a train system based on a multi-state complex network

A reliability evaluation of a rail transit train system is a core research technique to ensure the safety of train operation. However, the existing reliability analysis methods for rail transit train systems ignore the multiple states of the system. To overcome this shortcoming, as shown in Fig. 7, Ref. [16] proposes a novel rail transit train system reliability analysis method based on the improved d-minimal cut (d-MC). For example, the bogie system of a train has three functions: the bearing function, power transmission function, and buffering and damping function. Based on complex network theory, the function and interaction of the bogie system are analyzed. The reliability flow network of a train's bogie system is established by introducing the concept of load flow. Eliminating unnecessary candidate d-MCs and repeated d-MCs in the d-MC analysis solves the problem of having a large number of unnecessary candidate solutions and repeated solutions in the

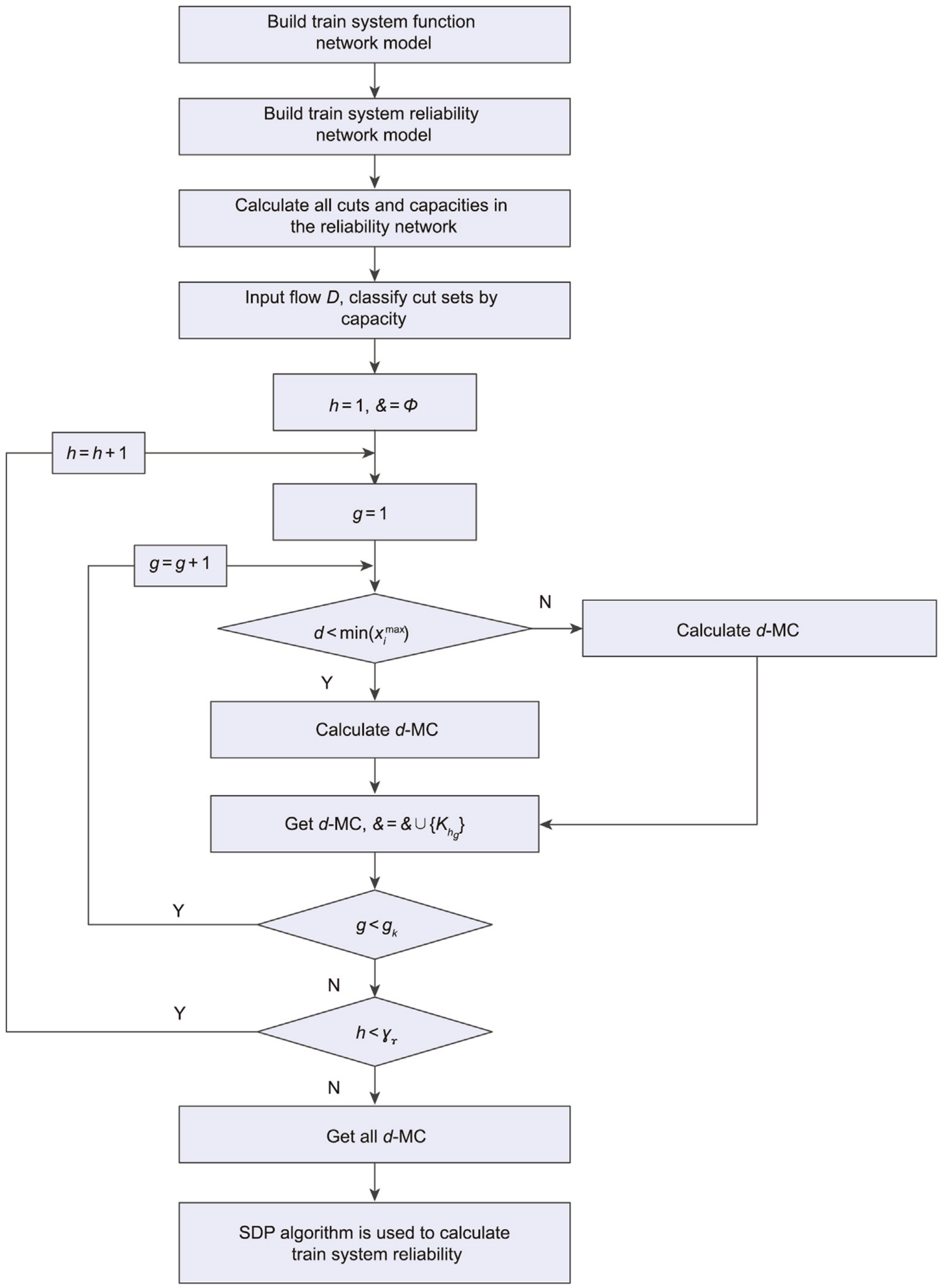

Fig. 7. A novel multi-state bogie system reliability analysis method based on an extended *d*-MC model for train systems. SDP: sum of disjoint products. The defination of all the abbreviations in this figure could be find in Ref. [16].

existing *d*-MC theory, and greatly improves the efficiency of the system reliability calculation.

The proposed method adds edge to and deforms the bogie system function relation network with multiple fountains and multiple confluxes to establish a unidirectional convergence network model. Then, it gives flow to the network edge in order to obtain the final bogie system reliability flow network model (Fig. S5(a) in Appendix A, dotted lines are additional edges, and the R and S nodes are virtual nodes). As shown in Fig. S5(b) in Appendix A, the improved *d*-MC model further improves the design reliability of the steering system under a certain load by adding edges. Table 2 shows that the reliability of the improved bogie system is slightly greater than that of the original bogie under various conditions, thereby demonstrating that the reliability of the rail transit train bogie system can be improved by increasing the mutual relations and interactions between components in the system.

4.2. Intelligent perception technology for the train operating environment

4.2.1. Railway video image enhancement based on a multiscale residual network

With the development of automatic train driving, the railway has higher requirements for safety in the surrounding environment. The intrusion of living things or objects onto the railway tracks is very harmful to the train operating environment, and

Table 2
Design reliability of the bogie system.

d	Bogie analysis results in Fig. S5(a)		Bogie analysis results in Fig. S5(b)	
	Number of cuts	R_d	Number of cuts	R_d
1	30	0.930831	34	0.949314
2	44	0.813449	46	0.815273
3	62	0.531583	82	0.534322
4	56	0.091839	98	0.091846

d: demand. R_d: reliability of demand d.

monitoring video has gradually become more common as an important means of detecting intrusion objects. The quality of railway surveillance video is significantly affected by the weather, especially bad weather such as haze, which seriously affects the accuracy of the object detection algorithm based on artificial intelligence. Therefore, it is necessary to develop an image enhancement algorithm to restore image quality affected by haze, rain, or other bad weather factors. Taking image haze removal (known as "image dehazing") as an example, there are two problems in railway monitoring image dehazing: image distortion and haze residuals.

To solve these two problems, Ref. [17] proposes a railway image dehazing algorithm based on a multiscale residual network. The algorithm process is shown in Fig. 8. First, multiscale convolution kernels are added based on the residual network, and more useful information can be extracted by integrating the features of three sets of convolution kernels. Next, a residual block, bottleneck layer, and loss function in a neural network are optimized to improve performance. In addition to optimizing the CNN, training datasets of railway scenes are synthesized. This method tests railway haze images and calculates the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), where the greater the PSNR and SSIM, the better. The results are shown in Fig. S6 in Appendix A. The proposed method achieves the best dehazing results compared to these classic image dehazing methods [25–30] in terms of both PSNR and SSIM, and the visual results are good without image distortion or haze residuals.

4.2.2. Railway small-object-intrusion detection based on machine vision

Railway intrusion seriously threatens railway safety and can cause huge numbers of casualties and enormous property losses. Video surveillance can be used to monitor railway sites in real time and has become the main method used to monitor the railway operation environment. However, the coverage distance of railway monitoring can range from several hundred meters to many kilometers, and such a long monitoring distance will result in the intrusion objects within images being extremely small in size.

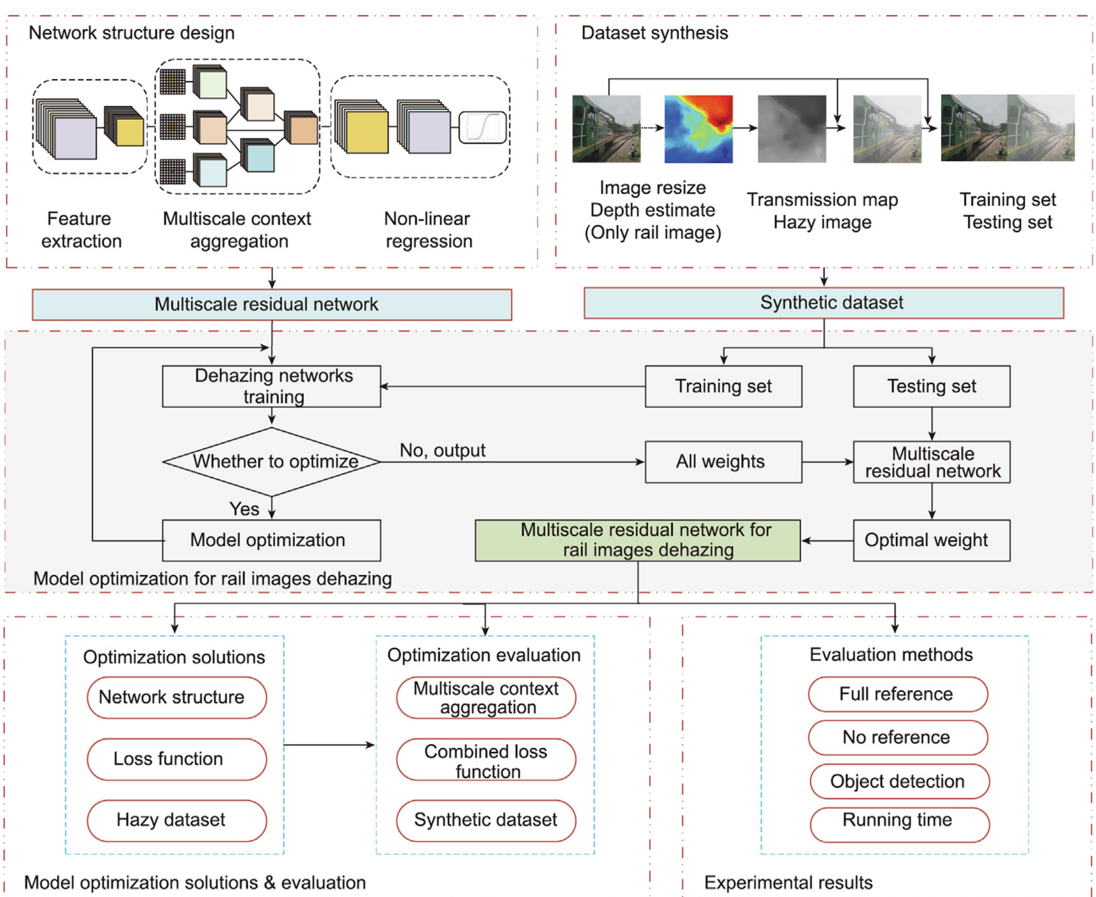

Fig. 8. Haze removal on a railway monitoring image based on a multiscale residual network.

Moreover, the texture and color characteristics of such small objects can be either not obvious or missing, creating further difficulty in detecting intrusion.

To improve the detection accuracy of the intelligent recognition algorithm for small intruding objects, Ref. [18] proposes a small-object intrusion detection method for railways, which is based on prior anchor clustering and the SE attention mechanism [19]. The model architecture is shown in Fig. 9. Since railway intrusion objects are generally small, *k*-means clustering is used to cluster the prior anchors of the intrusion object dataset. To improve the robustness and generalization of the algorithm, the railway datasets are amplified. The channel attention mechanism is used in the backbone network Darknet53 to enhance the interaction of spatial features on the channel. The test results are shown in Fig. S7 in Appendix A. The proposed method achieves good results; in comparison, YOLOv3 [20] and YOLOv3-SPP exhibit missing and false detection and are not good for small-object detection in railway scenes.

4.2.3. An intelligent inspection method for railway infrastructure using UAVs

The pantograph catenary system is a key part of the railway infrastructure that transmits power from substations to trains. However, the fasteners on catenary support devices (CSDs) often have defects, such as being loose, broken, or missing, due to the large contact force between the pantograph and the catenary. UAV inspection is an efficient, large-scale, and full-time inspection technology that can work together with comprehensive inspection vehicles and manual inspection to significantly improve the onsite inspection scope and work efficiency.

Hence, Ref. [21] proposes a UAV image-based component defect detection model with two improved CNNs: cascade YOLO (CYOLO) and rotation RetinaNet (RRNet). CYOLO is a cascaded YOLO network for locating CSD joints, and RRNet is a rotated box-based detector for inspecting arbitrary direction fasteners. As shown in Fig. S8 in Appendix A, CYOLO integrates shallow edge features with deep semantic information from images and only utilizes two scale prediction layers based on the joint proportion in images to improve location accuracy. Moreover, the Gridmask data-augmenting method is used to overcome the over-fitting problem caused by excessive similarity between data in the model training process. In the UAV images, the CSD fasteners have uncertain directions and are small in size, so it is difficult to detect fastener defects. As shown in Fig. S9 in Appendix A, the network structure and anchor box mechanism of RRNet can generate an appropriate number of candidate boxes of a suitable size for fastener inspection, thereby avoiding unnecessary computational overhead. Rotation intersection of union loss is introduced to overcome the angle imprecision of the predicted boxes. Figs. S10 and S11 in Appendix A show some examples of joint location and fastener inspection,

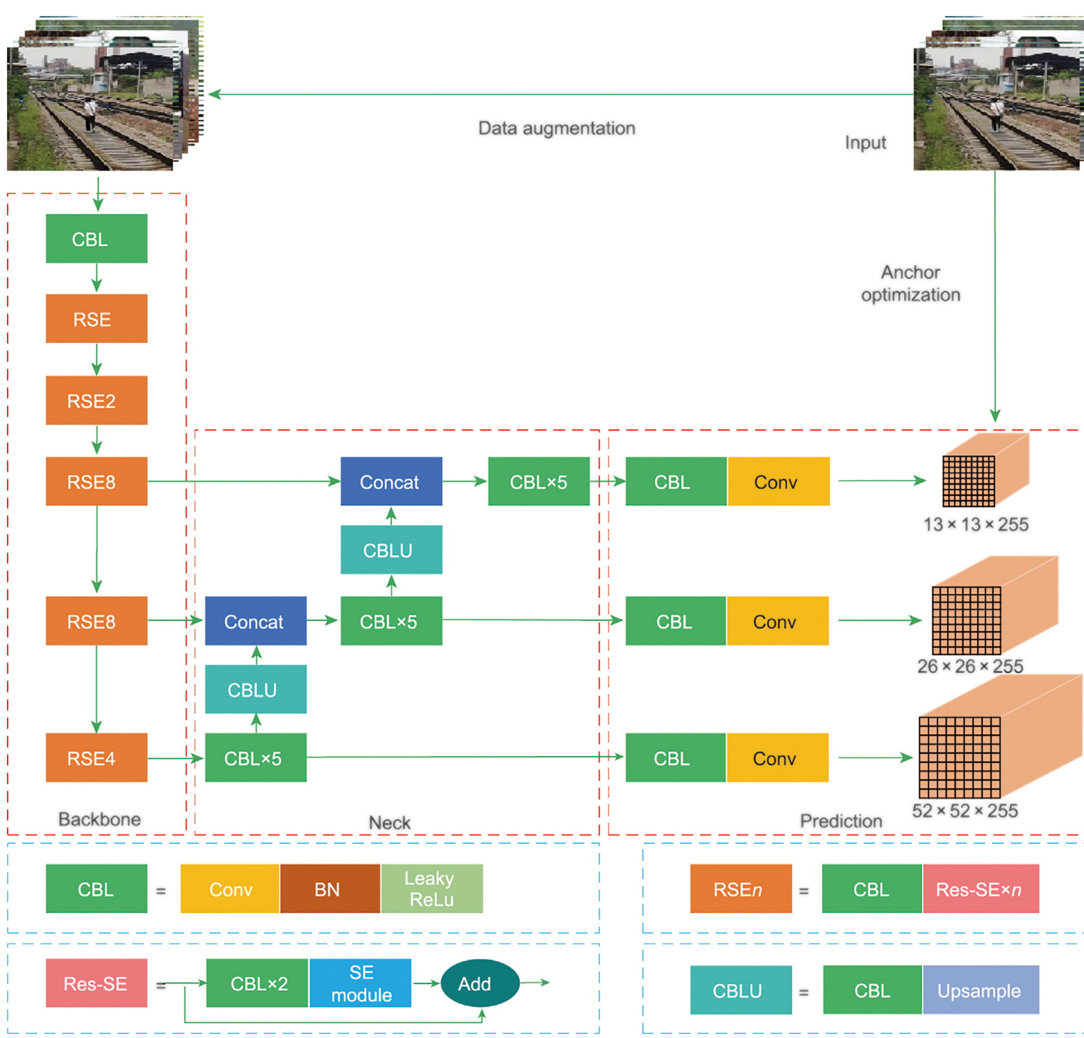

Fig. 9. A small-object-intrusion detection model for the railway perimeter. Conv: convolution; BN: BatchNorm. ReLu: Rectified Linear Unit; Concat: concatenate.

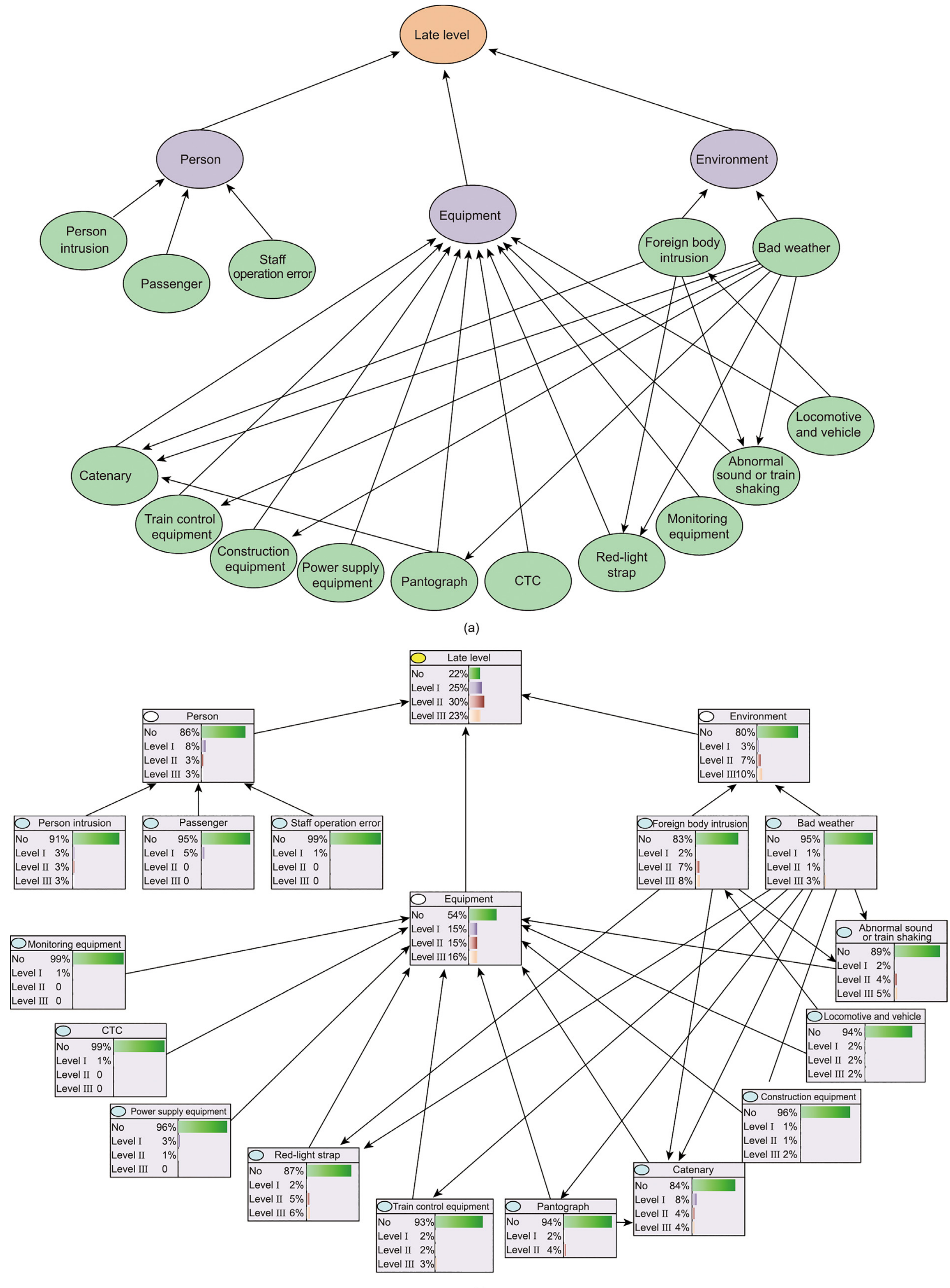

Fig. 10. Train running delay prediction based on a Bayesian network. (a) A Bayesian network structure of data learning modified according to expert experience; (b) a Bayesian network probability distribution of the initial delay of high-speed trains. CTC: centralized traffic control system.

respectively. As shown in Fig. S11, RRNet can effectively detect the arbitrary direction fastener on the different joints.

4.3. Railway network operation risk assessment and intelligent scheduling technology

4.3.1. Impact analysis of train delay during an emergency based on a random forest algorithm

The number of delayed trains caused by an emergency is an important evaluation index for measuring the impact of that emergency. After an emergency has occurred, if numerous trains are affected, then the effects of the emergency event are widespread, making it difficult for trains to resume normal operation. The greater the harm caused by an event is, the more attention should be paid to prevention.

To further calculate the degree of influence of different emergencies in terms of the number of delayed trains and to quantify the degree of harm caused by emergencies, Ref. [22] establishes a prediction model of delayed trains in an emergency based on a random forest algorithm. In the prediction model, the type, duration, place, and time of an emergency are taken as the characteristic variables, and the number of delayed trains caused by the emergency is taken as the object variable. The fitting degrees of the training and testing sets in the prediction model are 0.9801 and 0.9774, respectively, indicating that the model has effective results. The prediction results are shown in Fig. S12 in Appendix A; the predicted value of the model is basically consistent with the actual value.

4.3.2. Train running delay prediction based on a Bayesian network

Given the complexity of railway networks and the increase in driving density, risk factors such as equipment failure, bad weather, and human interference bring greater challenges to the punctual operation of high-speed trains. The reasons for an initial delay can be complex and interlocked, so it is necessary to study the correlation between initial delay and the reasons for the delay of high-speed trains. We establish an analysis method based on a Bayesian network that reflects the dependence between variables. First, an expert scoring method is used to score the correlation between the factors influencing the delay. Dempster–Shafer (D–S) evidence fusion theory is selected to reduce the subjectivity of the expert empirical judgment, so as to build a reasonable network structure. Then, as shown in Fig. 10 (a), a greedy thick thinning algorithm is adopted for structure learning in order to obtain the final network structure based on expert experience. Finally, the node parameters of the Bayesian network are calculated using an expectation maximization algorithm, and Bayesian inference is carried out. The calculation results are shown in Fig. 10 (b), where the bar graph of each node shows the edge probability of the corresponding variable. A joint tree algorithm is used to predict and deduce all causes of delay; the prediction results are summarized in Table 3.

Table 3
Prediction of train delays under different risk levels.

Types	Fault level		
	Level I fault	Level II fault	Level III fault
Staff operation error	No	Level I late	Level II late
Passenger	No	Level I late	Level II late
Person intrusion	No	Level I late	Level II late
Construction equipment fault	Level I late	Level II late	Level III late
Catenary fault	Level I late	Level II late	Level III late
Pantograph fault	Level I late	Level II late	–
Power supply equipment fault	Level I late	Level II late	Level III late
Monitoring equipment fault	Level I late	Level II late	Level III late
Locomotive and vehicle fault	Level I late	Level II late	Level III late
Train control fault	Level I late	Level II late	Level III late
Red-light strap fault	Level I late	Level II late	Level III late
Train shaking or abnormal sound	Level I late	Level II late	Level III late
CTC fault	No	Level II late	Level II late
Bad weather	Level I late	Level II late	Level II late
Foreign-body intrusion	Level I late	Level II late	Level II late

4.3.3. An intelligent rescheduling optimization method for high-speed trains in complex disruption scenarios

In an emergency situation, real-time train timetable rescheduling is an important tool for restoring railway operation order. At the macroscopic level, the high-speed railway timetable rescheduling problem is modeled as an event activity network composed of events and activities, and a mixed-integer linear programming model for train rescheduling optimization is proposed [23,24]. The proposed method uses numerous constraints, such as a train's order and the time between the end of a train's fault and the beginning of its operation; in some cases, the train's operation will be canceled. The method also includes other constraints related to the decreased capacity of the high-speed railway line. The objective functions and constraints are provided in Eqs. (3)–(15). In this model, rescheduling strategies such as train time adjustment, train operation sequence adjustment, and train outage are considered, and a hybrid heuristic evolutionary algorithm that combines a genetic algorithm and a particle swarm optimization algorithm is developed. The position vector and genetic evolution operator are reconstructed according to the departure time and arrival time of each train at the station. A case study of the Beijing–Shanghai high-speed railway in China was tested to verify the proposed model and algorithm. The train rescheduling scheme is shown in Fig. S13 in Appendix A. This intelligent method optimizes the train schedule for the high-speed railway and provides an excellent train rescheduling scheme for emergency scenarios. It also provides auxiliary decision-making for dispatchers in practice. Tables 4–6 respectively present the symbols, parameters, and decision variables of the proposed model.

The objective function is

$$\min \sum_{t \in T} \lambda_t y_t + \sum_{e \in E} \mu_e^+ d_e^+ + \sum_{e \in E^{\mathrm{arr}}} \mu_e^- d_e^- \tag{3}$$

The basic constraints are

$$2M_1 y_{t_e} - M_1 \le x_e - q_e \le M_1 \ \forall e \in E t_e \in T \tag{4}$$

$$D \ge d_e^+ \ge x_e - q_e - M_1 y_{t_e} \ \forall e \in E t_e \in T \tag{5}$$

$$d_e^- \ge q_e - x_e, d_e^- \ge D \ \forall e \in E^{\mathrm{arr}} \tag{6}$$

$$x_e \ge q_e \ \forall e \in E^{\mathrm{dep}} \tag{7}$$

Table 4
Model symbols.

Symbol	Description	Symbol	Description
A	Activity set	E^{arr}	Arrival event set
A_{head}	Head activity set	E^{dep}	Departure event set
$A_{\mathrm{head}}^{\mathrm{arr\text{-}arr}}$	Head activity set of the adjacent trains arrive at the station	$H_{\mathrm{dis}}^{\mathrm{start}}$	Segmentation fault occurrence time
$A_{\mathrm{head}}^{\mathrm{dep\text{-}dep}}$	Head activity set of the adjacent trains depart from the station	L_a	Minimum duration time of the activity
A_{station}	Station activity set	s	Station
A_{train}	Train activity set	S	Station set
a	Activity	T	Train set
C_s^{down}	Downward capability of station s	T^{down}	Downward train set
C_s^{up}	Upward capability of station s	T^{up}	Upward train set
E	Event set	t	Train
e	Event	t^{dep}	Departure time of train t from its departure station
f	The event before or after event e	t_e	Train corresponding to event e

Table 5
Model parameters.

Parameters	Description
D	Maximum allowable deviation time of the event
M_1	Positive integer, where M_1 is one cycle time: M_1 = 1440
M_2	Positive integer, $M_2 > M_1$
q_e	Indicates that event e occurs at the time given in the planned train diagram
u_e^+	Late penalty per unit time of event e
u_e^-	Early penalty per unit time of event e
λ_t	Penalty for the cancelation of train t operation

Table 6
Model decision variables.

Decision variables	Description
d_e^+	Late time of event e
d_e^-	Early time of event e
x_e	Time of event e in the adjusted train diagram
y_t	Indicates whether train t will be canceled, where 1 means that train t is canceled and 0 means the opposite
λ_{ef}	The sequence of events, where 1 means that event e occurred before event f, and 0 means the opposite
λ_{fe}	The sequence of events, where 1 means that event f occurred before event e, and 0 means the opposite
φ_a	Activities between two trains in the station; 0–1 variables are used to represent them

The constraints of single-train operation and stopping are

$$x_f - x_e \geq L_a \; \forall a \in (e,f) \in A_{\text{train}} \tag{8}$$

The constraints of the running interval of adjacent trains are

$$x_f - x_e + M_2(1-\lambda_{ef}) \geq L_a \; \forall a \in (e,f) \in A_{\text{head}} \tag{9}$$

$$\lambda_{ef} + \lambda_{fe} = 1 \; \forall (e,f) \in A_{\text{head}} \wedge (f,e) \in A_{\text{head}} \tag{10}$$

The capacity constraints are

$$\begin{aligned} &\sum_{a=(e,f)\in A_{\text{head}}^{\text{arr-arr}}} \lambda_{ef} - \sum_{a=(e,f)\in A_{\text{head}}^{\text{dep-dep}}} \varphi_a \leq C_s^{\text{down}} - 1 \\ &\quad \forall s \in S, f \in E_s^{\text{arr}} : q_f \geq H_{\text{dis}}^{\text{start}} \wedge t_f \in T^{\text{down}} \end{aligned} \tag{11}$$

$$\begin{aligned} &\sum_{a=(e,f)\in A_{\text{head}}^{\text{arr-arr}}} \lambda_{ef} - \sum_{a=(e,f)\in A_{\text{head}}^{\text{dep-dep}}} \varphi_a \leq C_s^{up} - 1 \\ &\quad \forall s \in S, f \in E_s^{\text{arr}} : q_f \geq H_{\text{dis}}^{\text{start}} \wedge t_f \in T^{\text{up}} \end{aligned} \tag{12}$$

$$x_f - x_e + M_2(1-\varphi_a) \geq L_a \; \forall a \in (e,f) \in A_{\text{station}} \tag{13}$$

The train cancelation constraints are

$$y_t = 0 \; \forall t \in Tt^{\text{dep}} \in H_{\text{dis}}^{\text{start}} \tag{14}$$

The constraints of train operation before the occurrence of disruptions are

$$x_e = q_e \; \forall e \in E : q_e \leq H_{\text{dis}}^{\text{start}} \tag{15}$$

5. Conclusions

The intelligent railway is a key development direction for the world's railway systems. As the central purpose of the intelligent railway, safety must be ensured and realized through continuous technological innovation. Active safety technology is the most effective way to address railway system security challenges. At present, breakthroughs have been made in the key technical fields of safety monitoring, intelligent inspection, fault diagnosis, predictive maintenance, operational risk prediction, and emergency dispatch command, and the effectiveness of the application of active safety technology in these fields has been verified. In the future, it will be necessary to strengthen research on non-destructive active sensing in areas such as non-destructive perception, self-powered technology, automated robots, and causal knowledge models, in order to develop technology and equipment systems and contribute to a global railway system without interruptions and death.

Acknowledgments

This paper is supported by the 2021 Chinese Academy of Engineering (CAE) International Top-level Forum on Engineering Science and Technology, "Safety and Governance of the High-Speed Railway."

Compliance with ethics guidelines

Yong Qin, Zhiwei Cao, Yongfu Sun, Linlin Kou, Xuejun Zhao, Yunpeng Wu, Qinghong Liu, Mingming Wang, and Limin Jia declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2022.06.025.

References

[1] State Council of the People's Republic of China. Ministry of Transport of the People's Republic of China: strive to move from a big transportation country to a strong transportation country [Internet]. Beijing: State Council of the People's Republic of China; 2017 Oct 18 [cited 2022 Mar 18]. Available from: http://www.gov.cn/zhuanti/2017-10/18/content_5232641.htm. Chinese.

[2] Qin Y, Sun X, Ma X, Wang M, Jia L, Tang T. Research on the architecture of intelligent railway 2.0 and its applications. J Beijing Jiaotong Univ 2019;43(1):138–45. Chinese.

[3] The European Rail Research Advisory Council. Rail route 2050: towards a competitive, resource efficient and intelligent rail transport system for 2050 [Internet]. Brussels: The European Rail Research Advisory Council; 2013 Jul 31 [cited 2022 Mar 18]. Available from: http://errac.org/publications/rail-route-2050-the-sustainable-backbone-of-the-single-european-transport-area/.

[4] The European Rail Research Advisory Council. Rail 2050 vision rail—the backbone of Europe's mobility. Report. Brussels: The European Rail Research Advisory Council; 2018 Jan 31.

[5] Europe's Rail. Shift2Rail [Internet]. Paris: Europe's Rail; 2015 Mar 31 [cited 2022 Mar 18]. Available from: https://shift2rail.org/research-development/.

[6] Department for Transport. Great British railways: the Williams–Shapps plan for rail. Report. London: Department for Transport; 2021 May 20.

[7] US Department of Transportation. Beyond traffic 2045. Report. Washington, DC: US Department of Transportation; 2015.

[8] IBM. Think beyond the rails: leading in 2025. Report. New York City: IBM Travel and Transportation Industry; 2016 May.

[9] Ministry of Land, Infrastructure, Transport and Tourism. Grand design of national spatial development towards 2050. Report. Trabzon: Ministry of Land, Infrastructure, Transport and Tourism; 2014 Apr 7.

[10] Railway Technical Research Institute. Master plan—research and development creating the future of railways—Research 2025. Report; 2019 Dec.

[11] China State Railway Group Co., Ltd. Outline of powerful nation railway advance planning in the new era. Report. Beijing: China State Railway Group Co., Ltd.; 2020 Aug 12. Chinese.

[12] Zhao X, Qin Y, He C, Jia L. Underdetermined blind source extraction of early vehicle bearing faults based on EMD and kernelized correlation maximization. J Intell Manuf 2022;33(1):185–201.

[13] Zhao X, Qin Y, He CB, et al. Intelligent fault identification for rolling element bearings in impulsive noise environments based on cyclic correntropy spectra and LSSVM. IEEE Access 2020;8:40925–38.

[14] Kou L, Qin Y, Zhao X, Chen X. A multi-dimension end-to-end CNN model for rotating devices fault diagnosis on high speed train bogie. IEEE Trans Vehicular Technol 2020;69(3):2513–24.

[15] Kou L. Service status identification and system reliability evaluation method based on tensor domain theory for rail train with monitoring data [dissertation]. Beijing: Beijing Jiaotong University; 2019. Chinese.

[16] Fu Y. System reliability analysis and maintenance optimization for rail transit train under complicated coupled interactions [dissertation]. Beijing: Beijing Jiaotong University; 2021. Chinese.

[17] Cao Z, Qin Y, Jia L, Xie Z, Liu Q, Ma X, et al. Haze removal of railway monitoring images using multi-scale residual network. IEEE Trans Intell Transp Syst 2021;22(12):7460–73.

[18] Liu Q. Intrusion image detection under typical severe weather conditions for high-speed railway—a deep learning-based method [dissertation]. Beijing: Beijing Jiaotong University; 2021. Chinese.

[19] Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018 Jun 18–22; Salt Lake City, UT, USA. Berlin: IEEE Xplore; 2018. p. 7132–41.

[20] Redmon J, Farhadi A. YOLOv3: an incremental improvement; 2018. arXiv:1804.02767.

[21] Wu Y, Qin Y, Qian Y, Guo F. Automatic detection of arbitrarily oriented fastener defect in high-speed railway. Autom Construct 2021;131:103913.

[22] Wang X. Impact analysis on high-speed railway operation emergency and research on the method for real-time timetable rescheduling [dissertation]. Beijing: Beijing Jiaotong University; 2020. Chinese.

[23] Wang M, Wang L, Xu X, Qin Y, Qin L. Genetic algorithm-based particle swarm optimization approach to reschedule high-speed railway timetables: a case study in China. J Adv Transport 2019;2019:6090742.

[24] Xu X. Optimization theory and method of high-speed train delay control under abnormal events. Beijing: China Communications Press; 2021.

[25] He K, Sun J, Tang X. Single image haze removal using dark channel prior. IEEE T Pattern Anal 2011;33(12):2341–53.

[26] Berman D, Treibitz T, Avidan S. Non-local image dehazing. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Berlin: IEEE Xplore; 2016. p. 1674–82.

[27] Ren W, Liu S, Zhang H, Pan J, Cao X, Yang M. Single image dehazing via multi-scale convolutional neural networks. In: Proceedings of the European conference on computer vision, 2016 Oct 8–16; Amsterdam, the Netherlands. Berlin: Springer; 2016. p. 154–69.

[28] Cai B, Xu X, Jia K, Qing C, Tao D. Dehazenet: an end-to-end system for single image haze removal. IEEE Trans Image Process 2016;25 (11):5187–98.

[29] Li B, Peng X, Wang Z, Xu J, Feng D. Aod-net: all-in-one dehazing network. In: Proceedings of the international conference on computer vision, 2017 Oct 22–29; Venice, Italy. Berlin: IEEE Xplore; 2017. p. 4780–8.

[30] Chen D, He M, Fan Q, Liao J, Zhang L, Hou D, et al. Gated context aggregation network for image dehazing and deraining. In: Proceedings of 2019 IEEE winter conference on applications of computer vision, 2019 Jan 7–11; Waikoloa, HI, USA. Berlin: IEEE Xplore; 2019. p. 1375–83.